its ability to code text impartially and reproducibly and its overall accuracy.

All diagnoses or clinical entries involving a particular area will generate exactly the same character in the area position of the code snippet. Thus, interrogating for injuries or events involving that area simply requires a search for the relevant code snippet embedded within the code. More importantly, the codes are grouped so that information can be derived from the placement and nature of the snippet contents. For example, all snippets with the letters A–Z at the area position are in a limb with upper case letters (A–Z) representing the leg and lower case letters (a–z) representing the arm (not shown). It should be noted that there is a correlation between the parts of the arm and leg, for example, the hip (A) corresponds to the shoulder (a).

To search for limbs, one simply searches for codes with letters at the area position. The areas adjacent to any area are one below and one above the ANSI representation of the current area. This allows meaningful groupings in interrogation. For example, the joints of any long bone will have snippets one below and one above the bone's own area snippet. For the bone of the thigh (the femur) the joints arc A (the hip) and C (the knee). These snippets can be further refined by modifiers like proximal, distal, medial lateral and anterior. Such modifiers have stereotyped snippets and during the process described they will be applied to the area snippet. "Lateral thigh" might be expressed as Bx and "Anterolateral thigh" as Bxy where x and y are snippets for lateral and anterior respectively.

Similar stereotyped snippets exist for bones, muscles, soft tissues and other parts of the body. For example, regarding the "femur" (bone of the thigh), if the snippet for "long bone" is "!" and the snippet for joint is "@" then "femur at the knee" might be expressed as "!B@C". This includes a surprising amount of information as we are not only discussing a bone but it is a long bone, associated with the thigh, and we are describing one of its joint surfaces, the distal one, which is at the knee joint. Interrogation involving any of these elements need only specify a pattern to which the codes must adhere. For example, injuries involving the knee joint all have "@C" at the defined position in the code while femur always starts with "!B" whether we are referring to the hip, the thigh or the knee.

Although the method of coding of the present invention has been described with reference to the coding of surgical data it is also applicable to other data using dictionaries appropriate to the application giving rise to the data. Also, it is obvious that the code snippets used in the example could be any characters or symbols other than those shown.

I claim:

1. A method of coding data representing diagnoses and/or operations which is carried out by a computer in which representative codes are assigned to said data comprising the steps of:

i) inputting said data to said computer,

ii) dividing said data into sub-parts, each said sub-part being of a particular sub-part type,

iii) preliminarily grouping said sub-parts into sub-groups by introducing divisions between some sub-parts in said data, each sub-part in a particular sub-group sharing an association with all other sub-parts in said particular sub-group,

iv) examining each said sub-group for completeness in regard to each sub-group containing a predetermined distribution of sub-part types,

v) if the examination of step (iv) reveals that the sub-group is not complete, completing said sub-groups by amending the constituent sub-parts so that each sub-group contains said predetermined distribution of sub-part types, and

vi) outputting said completed sub-groups as coded data for later interrogation or analysis.

2. A method of coding data as claimed in claim **1** wherein at least one list of records is provided and said data is a sequential stream of text and wherein said step of dividing said data into sub-parts comprises parsing said data whereby each word, or group of words, of said text is differentiated by comparing each said sub-part to records in said at least one list of records, said at least one list of records also indicating the type of records in said list.

3. A method of coding data as claimed in claim **2** wherein said step of dividing said data into sub-parts includes the step of associating the name of said list of records which contained the sub-part with the sub-part.

4. A method of coding data as claimed in claim **2** wherein said step of preliminarily grouping said sub-parts into sub-groups comprises the steps of:

i) comparing each said sub-part to records in said at least one list of records, said at least one list of records also indicating the type of records in the list, and

ii) representing each sub-part by a related record in said list or by an associated related record if the sub-part matches a record in one of said lists.

5. A method of coding data as claimed in claim **2** wherein said step of completing said sub-groups comprises the steps of:

i) carrying out a first examination of any sub-parts which may have a dual meaning in order to determine their appropriate meaning in the context of the surrounding sub-parts and inserting associated records from said at least one list of records to alter said sub-part or sub-group,

ii) examining the sub-groups adjacent each said sub-group in order to extract or copy sub-parts into an incomplete sub-group, and

iii) carrying out a second examination of sub-parts which may have a dual meaning to determine their appropriate meaning in the context of the surrounding sub-parts and inserting records from said at least one list of records to alter said sub-part or sub-group.

6. A method of coding data as claimed in claim **2** wherein at least one of said at least one list of records comprise terms relating to medical applications, at least one of said at least one list of records comprises anatomical terms relating to common names or descriptions of the bodily structure, at least one of said at least one list of records comprises clinical terms relating to the treatment of patients, and at least one of said at least one list of records comprises terms relating to the position of said anatomical terms.

7. A method of coding data as claimed in claim **1** wherein said divisions are introduced into said data in order to group associated sub-parts together and the positioning of said divisions is determined by one of a predetermined set of reasons and said step of preliminarily grouping said sub-parts also includes keeping a record of the reason for the positioning of divisions between sub-groups and identifying the list in which each sub-part was found.

8. A method of coding data as claimed in claim **1** wherein said divisions are introduced into said data in order to group associated sub-parts together and the positioning of said divisions is determined by one of a predetermined set of reasons and symbolic indicative codes represent said sub-parts, and wherein said symbolic representative codes are